On the Continuity of the Cramer Transform

Marianne Akian

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

On the continuity of the Cramer transform

M. Akian

N° 2841

Mars 1996

THÈME 4

Rapport de recherche

ISSN 0249-6399

On the continuity of the Cramer transform

M. Akian *

Thème 4 — Simulation
et optimisation
de systèmes complexes
Projet Méta2

Rapport de recherche n° 2841 — Mars 1996 — 19 pages

Abstract: The Cramer transform introduced in large deviations theory sends classical probabilities (resp. finite positive measures) into (min,+) probabilities (resp. finite measures) also called cost measures. We study its continuity when the two spaces of measures are endowed with the weak convergence topology. We prove that the Cramer transform is continuous in the subspace of logconcave measures and show counter examples in the opposite case. Moreover, in finite dimension, the Cramer transform is bicontinuous. Then, logconcave measures may be identified with lower semicontinuous convex functions.

Key-words: Max-plus algebra, Cramer transform, Logconcave measure, Logconcave function, weak convergence, Large deviations, Idempotent measure, Fenchel transform.

(Résumé : tsvp)

*e-mail : `Marianne.Akian@inria.fr`

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : (33 1) 39 63 55 11 – Télécopie : (33 1) 39 63 53 30

Sur la continuité de la transformée de Cramer

Résumé : La transformée de Cramer, introduite en théorie des grandes déviations, envoie les probabilités classiques (resp. les mesures positives finies) dans les probabilités (resp. les mesures finies) (min,+), appelées mesures de coût. Nous étudions sa continuité quand les espaces de mesures de départ et d'arrivée sont munis des topologies de la convergence faible. Nous prouvons que la transformée de Cramer est continue dans le sous espace des mesures logconcaves et donnons des contre-exemples dans le complémentaire. En dimension finie, la transformée de Cramer est de plus bicontinue. Les mesures logconcaves peuvent alors être identifiées à des fonctions convexes semi-continues inférieurement.

Mots-clé : Algèbre max-plus, Transformée de Cramer, Mesure logconcave, Fonction logconcave, Convergence faible, Grandes déviations, Mesure idempotente, Transformée de Fenchel.

Introduction

For any probability P on the Borel sets of a topological vector space E, the Cramer transform $\mathcal{C}(P)$ of P is defined by

$$\mathcal{C}(P) \stackrel{\text{def}}{=} \mathcal{F}(\log \mathcal{L}(P)).$$

Here, $\mathcal{L}$ and $\mathcal{F}$ denote respectively the Laplace and the Fenchel transforms, that is $\mathcal{L}(P) : \theta \in E' \mapsto \int_{x \in E} e^{\langle \theta, x \rangle} dP(x)$ and $\mathcal{F}(c) : x \in E \mapsto \sup_{\theta \in E'} \langle \theta, x \rangle - c(x)$ where E' denotes the dual space of E. Then $\mathcal{C}(P)$ is a lower semicontinuous (l.s.c.) convex function on E.

The Cramer transform was introduced in large deviations theory [11, 7, 14, 15, 27, 13]. Let (X_n) be a sequence of i.i.d. random variables with values in E and law P. If $\mathcal{L}(P)$ is finite in a neighborhood of zero, the Cramer theorem states that the law of $\frac{X_1 + \cdots + X_n}{n}$ obeys a large deviation principle with rate function $\mathcal{C}(P)$. This result is true if E is a finite dimensional vector space but also if it is locally convex and Hausdorff and if the support of P is a subset of a closed convex Polish subspace of E [8].

Another domain where the Cramer transform appears to be useful is decision theory. Following the theory of idempotent measures and integrals of Maslov [18], $(\min, +)$ probabilities (or finite $(\min, +)$ measures), that is probabilities in the idempotent semifield $(\mathbb{R} \cup \{+\infty\}, \min, +)$, can be introduced. In Polish spaces, they necessarily have a l.s.c. density [1] and the "measure of a set" corresponds to the minimum of a function (the density) on this set. Then, $(\min, +)$ probability theory leads to a probabilistic formalism for optimization and optimal control, that we call decision theory. This has been developed in [9, 12, 4, 2, 5]. Random variables (called decision variables) correspond to changes of variables or constraints parameters on an optimization problem. Markov chains (called Bellman chains) correspond to optimal control problems. Weak convergence of decision variables corresponds to the convergence of value functions. It is related with the epiconvergence introduced in convex analysis [6, 16] and is equivalent to the weak convergence of capacities introduced in [22, 21]. Classical limit theorems of probability (law of large numbers, central limit theorems) correspond to asymptotic results for the value function of optimal control problems.

Since the images of the Cramer transform are l.s.c. functions, we may say that the Cramer transform sends classical probabilities (resp. finite positive measures) into $(\min, +)$ probabilities (resp. finite $(\min, +)$ measures). For instance, if P is the Dirac measure at point m, $\mathcal{C}(P) = \mathbb{1}_m$ (where $\mathbb{1}_m(x) = +\infty$ for $x \neq m$ and $\mathbb{1}_m(m) = 0$), which is the density of the $(\min, +)$ Dirac measure. If P is the normal law $\mathcal{N}(m, \sigma)$ of mean m and standard deviation σ, $\mathcal{C}(P) = \mathcal{M}(m, \sigma) : x \mapsto \frac{1}{2} \left(\frac{x-m}{\sigma} \right)^2$ which has the same stability property as the Normal law. Moreover, the Cramer transform leads to another correspondence between probability theory and decision theory. Indeed, it sends the convolution of two probabilities into the $\inf$-convolution of their images; but the $(\min, +)$ equivalent of convolution is the inf-convolution of densities. Similarly, the Cramer transform sends independent random variables into independent decision variables.

In this paper we study the continuity of the Cramer transform when both classical probabilities and $(\min, +)$ probabilities spaces are endowed with the topology of the weak convergence. Logconcave measures and functions appear naturally in probability, heat equation theory, optimization [23, 19, 20, 10]. Typically, normal laws and Wiener measures are logconcave. We show here that the

Laplace transform has almost the same behavior as the Fenchel transform where convexity of functions is replaced by logconcavity of measures. We then prove the equivalence between vague and weak convergences on the subspace of logconcave finite measures over a finite dimensional vector space E and the bicontinuity of the Cramer transform on this subspace. Counter examples show that the logconcavity condition cannot be relaxed. For the infinite dimensional case, we prove a similar continuity result under assumptions on E similar to that of [8].

1 Notations and definitions

Let us first recall some definitions and results concerning $(\min, +)$ probabilities. The term $(\min, +)$ algebra refers to the idempotent semifield $(\mathbb{R} \cup \{+\infty\}, \min, +)$ denoted $\mathbb{R}_{\min}$. The neutral elements for the $\min$ and $+$ laws are respectively $+\infty$ and 0 and are denoted $\mathbb{0}$ and $\mathbb{1}$. Let us note that the order $\preceq$ associated with the idempotent $\min$ law is the reverse order $\geq$ of $\mathbb{R}$ (the order associated with an idempotent $\oplus$ law is defined by $a \preceq b \quad \Leftrightarrow \quad a \oplus b = b$). Therefore, inequalities and limits hold in the reverse order compared to that used for general idempotent or classical probabilities. The name $\mathbb{R}_{\min}$ will also be given to the set $\mathbb{R} \cup \{+\infty\}$ endowed with the topology defined by the order relation which is equivalent to that defined by the exponential distance $d(x, y) = |e^{-x} - e^{-y}|$. Let us finally note that, since $\mathbb{0} \in \mathbb{R}_{\min}$ is an upper bound of $\mathbb{R}_{\min}$ (for the classical order), bounded sets of $\mathbb{R}_{\min}$ correspond to lower bounded sets.

Definition 1. Let U be a topological space and $\mathcal{G}$ the set of its open sets. A finite $(\min, +)$ idempotent measure or cost measure on $(U, \mathcal{G})$ is an application $\mathbb{K}$ from $\mathcal{G}$ to $\mathbb{R}_{\min}$ such that

(a) $\mathbb{K}(\emptyset) = +\infty$

(b) $\mathbb{K}(\bigcup_n G_n) = \inf_n \mathbb{K}(G_n)$ for any $G_n \in \mathcal{G}$.

It is a $(\min, +)$ probability if in addition $\mathbb{K}(U) = 0$ and it is null if $\mathbb{K} \equiv +\infty$.

If c is a bounded function from U to $\mathbb{R}_{\min}$, $\mathbb{K}(G) = \inf_{u \in G} c(u)$ is a $(\min, +)$ idempotent measure. If $\mathbb{K}$ has this form, c is called a density of $\mathbb{K}$. Any cost measure $\mathbb{K}$ on $(U, \mathcal{G})$ admits a "maximal" (in terms of $\preceq$ order) extension $\mathbb{K}^*$ to the power set $\mathcal{P}(U)$ of U :

$$\mathbb{K}^*(A) = \sup_{G \supset A, G \in \mathcal{G}} \mathbb{K}(G).$$

If U is a separable metric space, then $\mathbb{K}$ has necessarily a density. Its "maximal" density is equal to $c^*(x) = \mathbb{K}^*(\{x\})$ and is l.s.c. [1, 17]. For general topological spaces, $\mathbb{K}^*(C) = \inf_{x \in C} c^*(x)$ for any compact set C [1]. Then $\mathbb{K}$ has a density if and only if $-\mathbb{K}^*$ is regular or is a capacity in the sense of [22], that is $\mathbb{K}(U) = \inf_{C \text{ compact} \subset U} \mathbb{K}^*(C)$.

In the sequel, 1_A and $\mathbb{1}_A$ denote respectively the classical and $(\min, +)$ characteristic functions of the set A : $1_A(x) = 1$ if $x \in A$ and 0 if $x \notin A$, $\mathbb{1}_A(x) = 0$ if $x \in A$ and $+\infty$ if $x \notin A$. If $A = \{x\}$, 1_A and $\mathbb{1}_A$ are denoted by 1_x and $\mathbb{1}_x$.

Given any cost measure $\mathbb{K}$ on $(U, \mathcal{G})$, the Maslov integral with respect to $\mathbb{K}$ is the unique $\mathbb{R}_{\min}$–linear form, denoted also $\mathbb{K}$, on the set of upper semicontinuous (u.s.c.) functions $f : U \to \mathbb{R}_{\min}$

such that $\mathbb{K}(f_n) \searrow \mathbb{K}(f)$ when $f_n \searrow f$ and $\mathbb{K}(\mathbb{1}_A) = \mathbb{K}(A)$ for $A \in \mathcal{U}$ [18, 1]. If the cost measure $\mathbb{K}$ has a density and c^* is its "maximal" density, then $\mathbb{K}(f) = \inf_{u \in U} f(u) + c^*(u)$. Therefore, the $(\min, +)$ equivalent to the Dirac measure δ_x in point x is the cost measure δ_x^c with density $\mathbb{1}_x$.

Using this formalism, the weak convergence of cost measures is defined as usual :

Definition 2. We say that $\mathbb{K}_n$ weakly converges towards $\mathbb{K}$, denoted $\mathbb{K}_n \stackrel{\mathrm{w}}{\to} \mathbb{K}$, if $\mathbb{K}_n(f) \to_n \mathbb{K}(f)$ for any bounded continuous function $f : U \to \mathbb{R}_{\min}$.

Using the correspondence between cost measures and their densities, we will also speak of the weak convergence of functions.

Theorem 3 ([5, Th. 5.2]). *Let $\mathbb{K}_n$ and $\mathbb{K}$ be cost measures on a metric space $(U, \mathcal{G})$. Then $\mathbb{K}_n \stackrel{\mathrm{w}}{\to} \mathbb{K}$ iff*

$$\begin{cases} \liminf_n \mathbb{K}_n(F) \geq \mathbb{K}(F) & \textit{for all closed } F \\ \limsup_n \mathbb{K}_n(G) \leq \mathbb{K}(G) & \textit{for all } G \in \mathcal{G}. \end{cases}$$

Let U be a locally compact topological space and suppose that $\mathbb{K}_n$ and $\mathbb{K}$ have densities (this is the case if U is a finite dimensional vector space). Then, they can be considered as capacities in the sense of [22] and the vague convergence may be defined either by the condition $\mathbb{K}_n(f) \to_n \mathbb{K}(f)$ for any continuous function f with compact support or by the following definition which coincides with that of [22].

Definition 4. We say that $\mathbb{K}_n$ vaguely converges towards $\mathbb{K}$, denoted $\mathbb{K}_n \stackrel{\mathrm{v}}{\to} \mathbb{K}$ if

$$\begin{cases} \liminf_n \mathbb{K}_n(C) \geq \mathbb{K}(C) & \text{for all compact } C \\ \limsup_n \mathbb{K}_n(G) \leq \mathbb{K}(G) & \text{for all } G \in \mathcal{G}. \end{cases}$$

The vague convergence is related to the epigraph convergence of functions defined in convex analysis [6, 16].

Definition 5. Let c_n and c be l.s.c. functions. We say that c_n converges in the epigraph sense (or epi-converges) towards c, denoted $c_n \stackrel{\mathrm{epi}}{\to} c$, if

(a) $\forall u,\ \forall u_n \underset{n}{\to} u,\ \liminf_n c_n(u_n) \geq c(u)$,

(b) $\forall u,\ \exists u_n \underset{n}{\to} u,\ \limsup_n c_n(u_n) \leq c(u)$.

Proposition 6 ([5, Th. 5.7]). *If U is a first countable topological space and $\mathbb{K}_n$ and $\mathbb{K}$ have l.s.c. densities c_n and c, then $\mathbb{K}_n \stackrel{\mathrm{v}}{\to} \mathbb{K}$ iff $c_n \stackrel{\mathrm{epi}}{\to} c$.*

The tightness is also defined in a similar way to classical probabilities (recall that $\mathbb{0} = +\infty$).

Definition 7. A set $\mathcal{K}$ of cost measures is tight iff

$$\sup_{C \text{ compact}} \inf_{\mathbb{K} \in \mathcal{K}} \mathbb{K}(C^c) = +\infty.$$

The vague and weak convergences are equivalent for tight sequences and any tight sequence of cost measures in a metric space admits a weakly convergent subsequence [5]. Let us also note that any sequence of l.s.c. functions on a separable metric space admits an epi-converging subsequence [6]. We also use the following result proved in convex analysis [6, 16]. A l.s.c. function $c : E \to \mathbb{R}_{\min}$ is proper iff $c \not\equiv \mathbb{0} = +\infty$.

Theorem 8. *Let E be a finite dimensional vector space. The Fenchel transform $\mathcal{F}$ is bicontinuous on the set $\mathcal{C}(E)$ of proper l.s.c. convex functions endowed with the topology of the epi-convergence.*

In a reflexive Banach space, the Fenchel transform is bicontinuous on $\mathcal{C}(E)$ endowed with the topology of the Mosco-epiconvergence defined as the epi-convergence (definition 5) except that the convergence of u_n towards u holds in the weak sense in (a) and in the strong sense in (b). The Mosco-epiconvergence is also equivalent to the vague convergence of cost measures defined as in definition 4 except that C may be any bounded closed convex set.

2 Statement of the results and counter examples

For any locally convex, Hausdorff topological vector space (l.c.h.t.v.s.) E, $\mathcal{LC}(E)$ will denote the set of non identically null u.s.c. logconcave functions, that is the set of functions $f : E \to \mathbb{R}^+$ such that $-\log f \in \mathcal{C}(E)$. If E is a closed convex subset of a l.c.h.t.v.s. X, any u.s.c. logconcave function f on E can be extended to an u.s.c. logconcave function on X by taking $f(x) = 0$ outside E. Then $\mathcal{C}(E) \simeq \{f \in \mathcal{C}(X),\ f(x) = 0\ \forall x \in E^c\}$. Generalizing the definition of [24] to the infinite dimensional case, we obtain

Definition 9. A positive measure μ defined on the Borel sets of a closed convex subset E of a l.c.h.t.v.s. X is said to be logarithmically concave (or logconcave) if

$$\mu(tA + (1-t)B) \geq \mu(A)^t \mu(B)^{1-t} \tag{1}$$

for any pair A, B of convex subsets of E and any $t \in [0, 1]$. The set of finite non null logconcave measures over E will be denoted by $\mathcal{CM}(E)$.

Again, $\mathcal{CM}(E) \simeq \{\mu \in \mathcal{CM}(X),\ \mu(E^c) = 0\}$.

Remark 10. From the definition of logconcave measures, we see that if θ is a continuous linear application between the l.c.h.t.v.s. E and F (in particular if $\theta \in E'$), the positive measure $\mu\theta^{-1}$ defined by $\mu\theta^{-1}(A) = \mu(\theta^{-1}(A))$ for any Borel set A of F is logconcave as soon as μ is logconcave. Indeed, if A and B are convex sets of F, $\theta^{-1}(A)$ and $\theta^{-1}(B)$ are convex and $\theta^{-1}(tA + (1-t)B) \supset t\theta^{-1}(A) + (1-t)\theta^{-1}(B)$. In particular, marginal laws of logconcave measures are also logconcave measures.

By theorems 1 and 2 of [24], any non null positive measure μ on a finite dimensional vector space E which admits a density $f \in \mathcal{LC}(E)$ with respect to the Lebesgue measure of E is logconcave. Moreover, as for logconcave measures, marginal laws of measures with logconcave density have logconcave density [24]. Indeed, logconcave measures have essentially a density. This is stated in the following result proved in appendix.

Proposition 11. *A finite non null positive measure μ on the finite dimensional vector space E is log-concave if and only if there exists an affine subspace F of E containing the support of μ and such that μ has a density $f \in \mathcal{LC}(F)$ with respect to the Lebesgue measure over F (if the dimension of F is zero, the Lebesgue measure is reduced to the Dirac measure and f is constant).*

Theorem 12. *Let E be a finite dimensional vector space. The Cramer transform is injective and bi-continuous from $\mathcal{CM}(E)$ to its image in $\mathcal{C}(E)$ endowed respectively with the classical weak convergence and the $(\min, +)$ weak convergence topologies. Moreover, for sequences in $\mathcal{CM}(E)$ vague convergence and weak convergence are equivalent.*

For the infinite dimensional case, we adopt the minimal assumptions on E stated by Bahadur and Zabell [8, 7, 13] for the construction of the Cramer transform and the Cramer theorem.

Theorem 13. *Let X be a l.c.h.t.v.s. and E be a closed convex subset of X such that E is a Polish space for the induced topology. The Cramer transform is continuous from $\mathcal{CM}(E)$ to $\mathcal{C}(E)$ endowed respectively with the classical weak convergence and the $(\min, +)$ weak convergence topologies.*

Remark 14. In [7] it is proved, using the Cramer theorem, that if $\operatorname{Supp}\mu \subset E$ closed convex, then $\operatorname{dom}\mathcal{C}(\mu) \subset E$. This can also be done directly using Hahn-Banach theorem. Indeed, if $x \notin E$, there exists $\theta_0 \in X$ such that $\langle \theta_0, x\rangle > 0$ and $\langle \theta_0, y\rangle \leq 0$ for any $y \in E$. Then $\mathcal{L}(\mu)(\lambda\theta_0) \leq 1$ for any $\lambda > 0$ and $\mathcal{C}(\mu)(x) \geq \sup_{\lambda > 0} \lambda\langle \theta_0, x\rangle = +\infty$. This implies that if E is a closed convex subset of X the image of $\mathcal{CM}(E)$ is included in $\mathcal{C}(E)$.

Example 15. Let us first give a counter example for the case of general positive measures or probabilities. Let P_n be the probability on $\mathbb{R}$ with density $p_n = e^{-c_n} / \int_{\mathbb{R}} e^{-c_n(x)}dx$ where

$$c_n(x) = \begin{cases} n\sqrt{x} & \text{for} \quad x \geq 0 \\ +\infty & \text{for} \quad x < 0. \end{cases}$$

Let us first note that c_n is a sequence of l.s.c. functions such that $c_n \xrightarrow{\mathrm{w}} \mathbb{1}_0$ for the $(\min, +)$ weak convergence, but that $\mathcal{F}(c_n) = \mathcal{F}(\mathbb{1}_{[0,+\infty)}) = \mathbb{1}_{(-\infty,0]}$ and then does not epi-converge towards $\mathcal{F}(\mathbb{1}_0) \equiv \mathbb{1}$. Thus c_n gives a counter example to the continuity of the Fenchel transform for general (not necessarily convex) proper l.s.c. functions.

We have also $P_n \xrightarrow{\mathrm{w}} \delta_0$ for the classical weak convergence, since

$$\int f(x)dP_n(x) = \frac{\int_0^{+\infty} f\left(\frac{x}{n^2}\right) e^{-\sqrt{x}}dx}{\int_0^{+\infty} e^{-\sqrt{x}}dx}$$

and

$$\left|\int f(x)dP_n(x) - f(0)\right| \leq \omega(f, \frac{1}{n}) + 2||f||_\infty \frac{\int_n^{+\infty} e^{-\sqrt{x}}dx}{\int_0^{+\infty} e^{-\sqrt{x}}dx},$$

where $\omega(f, .)$ denotes the continuity modulus of f in 0.

However

$$\mathcal{L}(P_n)(\theta) = \mathcal{L}(P_1)\left(\frac{\theta}{n^2}\right) \begin{cases} = +\infty \text{ for } \theta > 0, \\ \xrightarrow[n]{} \mathcal{L}(P_1)(0) = 1 \text{ for } \theta \leq 0. \end{cases}$$

Then similarly to $\mathcal{F}(c_n)$ we have $\log\mathcal{L}(P_n) \to_n \mathbb{1}_{(-\infty,0]}$ pointwise and in the epigraph sense. Then $\mathcal{C}(P_n) \overset{\text{epi}}{\to} \mathbb{1}_{[0,+\infty)} \neq \mathcal{C}(\delta_0) = \mathbb{1}_0$.

This example shows in addition the analogy of behavior between the Fenchel transform and the logarithm of the Laplace transform.

Example 16. Let us give another counter example concerning stable distributions. Let us consider a stable distribution μ with order $0 < \alpha < 2$ on $\mathbb{R}$. Then, either the domain of $\mathcal{L}(\mu)$ is $\{0\}$ so the Cramer transform of μ has no interest. Either it is equal to one half space, for instance $\mathbb{R}^+$ and then $\log\mathcal{L}(\mu)(\theta) = \lambda\theta^\alpha$ for $\theta \geq 0$ and $\log\mathcal{L}(\mu)(\theta) = +\infty$ for $\theta < 0$. Considering now the probability $\mu_n = \mu(n.)$, we get $\mu_n \overset{\text{w}}{\to} \delta_0$. However $\log\mathcal{L}(\mu_n)(\theta) = \log\mathcal{L}(\mu)\left(\frac{\theta}{n}\right) \to_n \mathbb{1}_{[0,+\infty)}(\theta)$ for any θ and $\mathcal{C}(\mu_n) \overset{\text{epi}}{\to} \mathbb{1}_{(-\infty,0]} \neq \mathcal{C}(\delta_0) = \mathbb{1}_0$. This shows that the law of large numbers for i.i.d. random variables with a stable distribution of order $\alpha \neq 2$ cannot be transferred from probability to optimization by the Cramer transform. Indeed, the $(\min,+)$ law of large numbers does not work in general for independent identically costed decision variables with non tight cost density such as $\mathcal{C}(\mu) = \lambda' \max(0,-x)^{\alpha'}$ (with $\lambda' > 0$ and $\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$).

Remark 17. In [2], we noticed that if the Cramer transform were continuous, the $(\min,+)$ law of large numbers [25, 4, 2, 5, 12] and the $(\min,+)$ central limit theorem [25, 4, 5] may have been consequences of the classical ones. Example 15 and 16 show that this is not the case. Indeed, the laws appearing in these results are in general not the images of probabilities by the Cramer transform. By its definition $\mathcal{C}(P)$ is necessarily a proper l.s.c. convex function. Moreover, if for instance $E = \mathbb{R}$, $\mathcal{F}(\mathcal{C}(P)) = \log\mathcal{L}(P)$ is analytical and then infinitely differentiable in the interior of its domain. Then, the l.s.c. function $\mathcal{M}^p(m,\sigma) : x \mapsto \frac{1}{p}\left|\frac{x-m}{\sigma}\right|^p$ with $p > 1$, $p \neq 2$ and $\sigma > 0$ appearing in central limit theorem [4, 5] is not the image of a probability by the Cramer transform. Indeed, $\mathcal{F}(\mathcal{M}^p(m,\sigma))(\theta) = m\theta + \frac{1}{p'}|\sigma\theta|^{p'}$ with $\frac{1}{p} + \frac{1}{p'} = 1$, then it is not regular in 0 except for integer values of p'. But in that last case $p' > 2$ and if $\mathcal{C}(P_X) = \mathcal{M}^p(m,\sigma)$, $E(X) = m$ and the variance of X is zero, then $X \equiv m$ and $\sigma = 0$.

Plan of the proof of Theorems 12 and 13. Let $\mathcal{M}(E)$ denotes the set of finite positive measures on the Borel sets of E. If E is a closed subset of X, $\mathcal{M}(E) \simeq \{\mu \in \mathcal{M}(X),\ \mu(E^c) = 0\}$. If E is a Polish space, the topology of the classical weak convergence is metrizable. Then the continuity of the Cramer transform from a subset of $\mathcal{M}(E)$ to $\mathcal{C}(E)$ both endowed with the weak convergence topologies is equivalent to the sequential continuity that is $\mu_n \overset{\text{w}}{\to} \mu \Rightarrow \mathcal{C}(\mu_n) \overset{\text{w}}{\to} \mathcal{C}(\mu)$. Since the topology of the $(\min,+)$ weak convergence is also metrizable in the subspace of tight $(\min,+)$ measures [3], the bicontinuity of the Cramer transform on $\mathcal{CM}(E)$ is equivalent to the sequential bicontinuity, if $\mathcal{C}(\mu)$ is proved to be tight for any $\mu \in \mathcal{CM}(E)$ which is done in section 4 and 5.

Since, in a reflexive Banach space, the $(\min,+)$ weak convergence implies the Mosco-epiconvergence and the Fenchel transform is bicontinuous on l.s.c. proper convex functions for the Mosco-epiconvergence, the continuity of $\log\mathcal{L}$ is required at least in the following sense :

$$\mu_n \overset{\text{w}}{\to} \mu \Rightarrow \log\mathcal{L}(\mu_n) \overset{\text{epi}}{\to} \log\mathcal{L}(\mu). \tag{2}$$

However the continuity of $\log \mathcal{L}$ for the weak convergence is not required. Indeed if $x_n \to_n 0$, $\delta_{x_n} \stackrel{\text{w}}{\to} \delta_0$, $\mathcal{C}(\delta_{x_n}) = \mathbb{1}_{x_n} \stackrel{\text{w}}{\to} \mathcal{C}(\delta_0) = \mathbb{1}_0$ and $\log \mathcal{L}(\delta_{x_n}) = \mathcal{F}(\mathbb{1}_{x_n}) : \theta \mapsto \theta x_n \stackrel{\text{epi}}{\to} \log \mathcal{L}(\delta_0) \equiv 0$ but $\log \mathcal{L}(\delta_{x_n})$ does not weakly converge towards 0.

If (2) holds in finite dimension, the continuity of the Cramer transform will be a consequence of the tightness of $\mathcal{C}(\mu_n)$ for tight sequences μ_n [5]. In the infinite dimensional case, we apply Theorem 12 to marginal laws. Then $\mu_n \stackrel{\text{w}}{\to} \mu$ with μ_n and μ in $\mathcal{CM}(E)$ implies $\mathcal{C}(\mu_n)\theta^{-1} \stackrel{\text{w}}{\to} \mathcal{C}(\mu)\theta^{-1}$ in $\mathcal{C}(\mathbb{R})$ (since $\mu_n\theta^{-1} \stackrel{\text{w}}{\to} \mu\theta^{-1}$ and $\mathcal{C}(\mu)\theta^{-1} = \mathcal{C}(\mu\theta^{-1})$) for any $\theta \in X'$. If $\mathcal{C}(\mu_n)$ is tight, there exists a converging subsequence also denoted $\mathcal{C}(\mu_n) \stackrel{\text{w}}{\to} \varphi$, where φ is a l.s.c. function on E. By the convexity of $\mathcal{C}(\mu_n)$, φ is convex and can be extended to a l.s.c. convex function on X. Then, $\mathcal{C}(\mu_n)\theta^{-1} \stackrel{\text{w}}{\to} \varphi\theta^{-1} = \mathcal{C}(\mu)\theta^{-1}$ for any $\theta \in X'$ and $\mathcal{C}(\mu) = \varphi$. By unicity of the limit, the sequence $\mathcal{C}(\mu_n)$ weakly converges towards $\mathcal{C}(\mu)$. Again the continuity of the Cramer transform is a consequence of the tightness of $\mathcal{C}(\mu_n)$ for tight sequences μ_n.

We then reduce the proof of Theorems 12 and 13 to three parts : the continuity of $\log \mathcal{L}$ for the epiconvergence at least in finite dimension (section 3), the tightness of $\mathcal{C}(\mu_n)$ in finite dimension (section 4) and then in infinite dimension (section 5). The tightness of $\mathcal{C}(\mu)$ for any $\mu \in \mathcal{CM}(E)$, the injectivity and the continuity of $\mathcal{C}^{-1}$ in finite dimension are proved in section 4.

3 Continuity of $\log \mathcal{L}$

We use the following lemma which is the analogue of a result proved in [2] for the Fenchel transform. The vague and weak convergences are denoted identically for classical and $(\min, +)$ measures.

Lemma 18. *Let E and X be as in Theorem 13 and X' be endowed with a topology such that $(x, \theta) \in X \times X' \mapsto \langle \theta, x \rangle$ is continuous. Let μ_n and $\mu \in \mathcal{M}(E)$. Then $\mu_n \stackrel{\text{w}}{\to} \mu$ implies*

$$\liminf_n \mathcal{L}(\mu_n)(\theta_n) \geq \mathcal{L}(\mu)(\theta) \quad \forall \theta_n \underset{n}{\to} \theta \text{ in } X'. \tag{3}$$

and

$$\mathcal{L}(\mu_n)(\theta) \underset{n}{\to} \mathcal{L}(\mu)(\theta) \quad \forall \theta,\ \exists t > 1,\ \limsup_n \mathcal{L}(\mu_n)(t\theta) < +\infty.$$

Proof. Let C be a compact subset of E, d a distance uniformly equivalent to the topology of E and $f(x) = \max(1 - \frac{d(x,C)}{\varepsilon}, 0)$. Then, $e^{\langle \theta, x \rangle} f(x)$ is bounded and continuous. Since $\mu_n \stackrel{\text{w}}{\to} \mu$ and

$$\mathcal{L}(\mu_n)(\theta) \geq \int e^{\langle \theta, x \rangle} f(x) d\mu_n(x),$$

we have

$$\liminf_n \mathcal{L}(\mu_n)(\theta) \geq \int_C e^{\langle \theta, x \rangle} d\mu(x).$$

Since E is a Polish space, μ is regular (or is a capacity), then taking the supremum over the compact set C, we get (3) when $\theta_n \equiv \theta$. For the case $\theta_n \to_n \theta$ we obtain

$$\liminf_n \mathcal{L}(\mu_n)(\theta_n) \geq e^{-\limsup_n g(C,n)} \int_C e^{\langle \theta, x \rangle} d\mu(x)$$

where

$$\begin{aligned} g(C,n) &= \sup_{d(x,C)<\varepsilon} |\langle \theta_n - \theta, x\rangle| \\ &\leq \sup_{x\in C} |\langle \theta_n - \theta, x\rangle| + \sup_{d(x,y)<\varepsilon} |\langle \theta_n - \theta, y-x\rangle| . \end{aligned}$$

Since the first term tends to 0 and the second is small when n is large and ε is small, we obtain (3).

For the second point, we use that $\mu_n(f) \to_n \mu(f)$ not only for bounded continuous functions f but also for functions f continuous and equi-integrable with respect to μ_n, that is such that

$$\int_{|f(x)|\geq k} f(x) d\mu_n(x) \underset{k\to\infty}{\to} 0$$

uniformly in n.

Since $\limsup_n \mathcal{L}(\mu_n)(t\theta) < +\infty$, we have $\mathcal{L}(\mu_n)(t\theta) \leq C$ for n large enough and

$$\int_{\langle\theta,x\rangle\geq k} e^{\langle\theta,x\rangle} d\mu_n(x) \leq \mathcal{L}(\mu_n)(t\theta) e^{(1-t)k} \underset{k\to\infty}{\to} 0$$

uniformly in n. □

Remark 19. In finite dimension, the proof of inequality (3) only requires the vague convergence of μ_n towards μ that is $\mu_n(f) \to_n \mu(f)$ for any continuous function with compact support.

Let us first give a proof of the continuity of the Fenchel transform that can be adapted to prove the continuity of $\log \mathcal{L}$. If the cost measure $\mathbb{K}$ has a l.s.c. density c, we write $\mathcal{F}(\mathbb{K})$ for $\mathcal{F}(c)$.

Proposition 20. *Let E, X and X' be as in Lemma 18. Let $\mathbb{K}_n$ be a sequence of cost measures on E with convex l.s.c. densities, then $\mathbb{K}_n \overset{\mathrm{w}}{\to} \mathbb{K} \Rightarrow \mathcal{F}(\mathbb{K}_n) \overset{\mathrm{epi}}{\to} \mathcal{F}(\mathbb{K})$ in X'.*

Proof. By a result equivalent to Lemma 18 for $\mathcal{F}$ instead of $\mathcal{L}$ [2], we have for any sequence $\theta_n \to_n \theta$

$$\liminf_n \mathcal{F}(\mathbb{K}_n)(\theta_n) \geq \mathcal{F}(\mathbb{K})(\theta)$$

that is condition (a) of definition 5 is fulfilled. Moreover $\mathcal{F}(\mathbb{K}_n)(0) \to_n \mathcal{F}(\mathbb{K})(0)$ since the constant function 0 is bounded.

Let $\theta \in X'$ such that $t\theta \in \operatorname{dom} \mathcal{F}(\mathbb{K})$ for some $t > 1$ and let prove that $\limsup_n \mathcal{F}(\mathbb{K}_n)(\theta) \leq \mathcal{F}(\mathbb{K})(\theta)$. We denote by c and c_n the l.s.c. densities of $\mathbb{K}\theta^{-1}$ and $\mathbb{K}_n\theta^{-1}$ ($\mathbb{K}\theta^{-1}$ is the cost measure on $\mathbb{R}$ such that $\mathbb{K}\theta^{-1}(A) = \mathbb{K}(\theta^{-1}(A))$ for any open set A of $\mathbb{R}$). The convexity of the density of $\mathbb{K}_n$ implies that of c_n. Then

$$\mathcal{F}(\mathbb{K})(t\theta) = \sup_{x\in\mathbb{R}} tx - c(x) < +\infty$$

implies that $x - c(x)$ decreases in some point that is $x_1 - c(x_1) > x_2 - c(x_2)$ for $x_1 < x_2$. Since $\mathbb{K}_n\theta^{-1} \overset{\mathrm{w}}{\to} \mathbb{K}\theta^{-1}$, $c_n \overset{\mathrm{epi}}{\to} c$ and there exists $x_n \to_n x_1$ such that $x_n - c_n(x_n) > x_2 - c_n(x_2)$ and

$x_n < x_2$ for n large enough. By the convexity of c_n, this implies that $x - c_n(x)$ decreases for $x \geq x_2$ and

$$\mathcal{F}(\mathbb{K}_n)(\theta) = \sup_x x - c_n(x) = \sup_{x \leq x_2} x - c_n(x).$$

Then

$$\limsup_n \mathcal{F}(\mathbb{K}_n)(\theta) \leq \sup_{x \leq x_2 + 1} x - c(x) \leq \mathcal{F}(\mathbb{K})(\theta).$$

This proves that $\limsup_n \mathcal{F}(\mathbb{K}_n)(\theta) \leq \mathcal{F}(\mathbb{K})(\theta)$ for θ such that $t\theta \in \operatorname{dom} \mathcal{F}(\mathbb{K})$ with $t > 1$. Using the left continuity of $t \mapsto \mathcal{F}(\mathbb{K})(t\theta)$ in $t = 1$, we can construct a sequence $t_n \to_n 1_-$ such that $\limsup_n \mathcal{F}(\mathbb{K}_n)(t_n\theta) \leq \mathcal{F}(\mathbb{K})(\theta)$ for any $\theta \in \operatorname{dom} \mathcal{F}(\mathbb{K})$ and then for any $\theta \in X'$ which proves property (b) of definition 5. $\square$

Let us now give the proof of the continuity of $\log \mathcal{L}$ which is a direct adaptation of the previous proof.

Proposition 21. *Let E, X and X' be as in Lemma 18 and let μ_n and $\mu \in \mathcal{CM}(E)$. Then*

$$\mu_n \stackrel{\text{w}}{\to} \mu \Rightarrow \log \mathcal{L}(\mu_n) \stackrel{\text{epi}}{\to} \log \mathcal{L}(\mu) \quad \textit{in } X'.$$

Proof. Again, by Lemma 18, $\liminf_n \log \mathcal{L}(\mu_n)(\theta_n) \geq \log \mathcal{L}(\mu)(\theta)$ for any $\theta_n \to_n \theta$. Let us prove that $\limsup_n \mathcal{L}(\mu_n)(\theta) \leq \mathcal{L}(\mu)(\theta)$ for any $\theta \in \operatorname{dom} \mathcal{L}(\mu)$. Since $\mu_n\theta^{-1}$ and $\mu\theta^{-1}$ are logconcave, $\mathcal{L}(\mu)(\theta) = \mathcal{L}(\mu\theta^{-1})(1)$ and $\mu_n\theta^{-1} \stackrel{\text{w}}{\to} \mu\theta^{-1}$, we can suppose that $E = \mathbb{R}$. In that case, μ_n is a Dirac measure or has logconcave density, then $e^{\theta x}\mu_n$ are not necessarily finite logconcave measures such that $e^{\theta x}\mu_n$ vaguely converges towards the finite logconcave measure $e^{\theta x}\mu$. We are then reduced to prove that $\mu_n \stackrel{\text{v}}{\to} \mu$ with μ_n not necessarily finite logconcave Radon measures and μ a finite logconcave measure over $\mathbb{R}$ imply $\limsup_n \mu_n(\mathbb{R}) \leq \mu(\mathbb{R})$.

Since $0 < \mu(\mathbb{R}) < +\infty$, we can suppose after translation that $\mu([0,1)) > 0$. Then, since

$$\mu(\mathbb{R}) = \sum_{k \in \mathbb{Z}} \mu([k, k+1)) < +\infty,$$

$k \mapsto \mu([k, k+1))$ decreases in some point $k = l \geq 1$, that is

$$\mu([l, l+1)) < e^{-\varepsilon}\mu([l-1, l)) \tag{4}$$

with $\varepsilon > 0$.

If μ has a density then the sets $[k, k+1)$ are μ-continuous for any k. Otherwise, from the logconcavity of μ, μ is a Dirac measure at a point $x \in [0,1)$. After translation, we can suppose that $x \in (0,1)$ and then the sets $[k, k+1)$ are again μ-continuous for any integer k. Therefore, if $\mu_n \stackrel{\text{v}}{\to} \mu$, (4) is valid for μ_n instead of μ and n large enough.

Since μ_n is logconcave,

$$\mu_n([l, l+1))^2 \geq \mu_n([l-1, l))\mu_n([l+1, l+2))$$

and applying (4) we obtain $\mu_n([l+1,l+2)) \leq e^{-\varepsilon}\mu_n([l,l+1))$ and so on. Then for $k \geq 0$ and $p \geq 1$,

$$\mu_n([(k+l)p,(k+l+1)p)) \leq e^{-\varepsilon kp}\mu_n([lp,(l+1)p))$$

and

$$\begin{aligned}\mu_n([0,+\infty)) &= \mu_n([0,lp)) + \sum_{k\in\mathbb{N}} \mu_n([(k+l)p,(k+l+1)p)) \\ &\leq \frac{1}{1-e^{-\varepsilon p}}\mu_n([0,(l+1)p)). \end{aligned} \tag{5}$$

Increasing p and l and decreasing ε, we obtain an analogue inequality for $\mu_n((-\infty,0])$. Then,

$$\limsup_n \mu_n(\mathbb{R}) \leq \frac{1}{1-e^{-\varepsilon p}}\mu([-(l+1)p,(l+1)p]) \leq \frac{\mu(\mathbb{R})}{1-e^{-\varepsilon p}}.$$

Taking the limit when p goes to infinity, we obtain the expected inequality. □

From the previous proof and Remark 19, we see that if $E = \mathbb{R}$, the vague convergence of μ_n towards μ is only required. This fact is also true in finite dimension and indeed allows to prove the equivalence between weak and vague convergences.

Proposition 22. *Let μ_n and $\mu \in \mathcal{CM}(\mathbb{R}^d)$, then*

$$\mu_n \xrightarrow{\text{v}} \mu \Leftrightarrow \mu_n \xrightarrow{\text{w}} \mu$$

Proof. The vague convergence of general finite measures on $\mathbb{R}^d$ is equivalent to the conditions

$$\begin{cases} \limsup_n \mu_n(C) \leq \mu(C) & \text{for all compact } C \\ \liminf_n \mu_n(G) \geq \mu(G) & \text{for all open } G. \end{cases}$$

and the weak convergence is equivalent to the same conditions but with closed sets in place of compact sets. Since the inequality on open sets implies $\limsup_n \mu_n(F) \leq \limsup_n \mu_n(\mathbb{R}^d) - \mu(F^c)$ for any closed sets, we can pass from vague to weak convergence only by proving $\limsup_n \mu_n(\mathbb{R}^d) \leq \mu(\mathbb{R}^d)$ or equivalently $\lim_n \mu_n(\mathbb{R}^d) = \mu(\mathbb{R}^d)$. This property has been proved in dimension $d = 1$ in the proof of previous proposition and then Proposition 22 is true in that case. Let us prove this property by induction on d. We suppose that it is true for $d-1$ and consider μ_n and $\mu \in \mathcal{CM}(\mathbb{R}^d)$ such that $\mu_n \xrightarrow{\text{v}} \mu$. Let C be a bounded convex set of $\mathbb{R}$ such that $\mu(\mathbb{R}^{d-1} \times \partial C) = 0$ and $\mu(\mathbb{R}^{d-1} \times C) > 0$ and consider the following measures on $\mathbb{R}^{d-1}$: $\mu_n^C(A) = \mu_n(A \times C)$ and $\mu^C(A) = \mu(A \times C)$. These measures are logconcave and finite, μ^C is non null and since $\liminf_n \mu_n^C(\mathbb{R}^{d-1}) \geq \mu(\mathbb{R}^{d-1} \times \operatorname{int} C) > 0$, μ_n^C is non null for n large enough. Moreover $\mu_n^C \xrightarrow{\text{v}} \mu^C$. By induction, we get $\mu_n(\mathbb{R}^{d-1} \times C) \to_n \mu(\mathbb{R}^{d-1} \times C)$ for any bounded convex set C such that $\mathbb{R}^{d-1} \times C$ is μ-continuous with positive measure. The positivity condition can be eliminated by adding a disjoint convex set C' satisfying the previous conditions. Considering now the measures $\mu_n'(A) = \mu_n(\mathbb{R}^{d-1} \times A)$ and $\mu'(A) = \mu(\mathbb{R}^{d-1} \times A)$ on $\mathbb{R}$ and noting that the proof of Proposition 21 only used the convergence of $\mu_n'(C)$ towards $\mu'(C)$ for μ'-continuous bounded convex sets C, we conclude that $\mu_n(\mathbb{R}^d) \to_n \mu(\mathbb{R}^d)$. □

4 Tightness of $\mathcal{C}(\mu_n)$ and bicontinuity of the Cramer transform in finite dimension

We use the following lemmas.

Lemma 23. *If E has finite dimension and $||\cdot||$ denotes a norm on E, the following propositions are equivalent for any finite positive measure μ on E*

(a) $0 \in \operatorname{int}\operatorname{dom}\mathcal{L}(\mu)$

(b) there exists $\varepsilon > 0$ such that $\int e^{\varepsilon\|x\|} d\mu(x) < +\infty$

(c) $0 \in \operatorname{int}\operatorname{dom}\mathcal{L}(\mu\theta^{-1})$ for any $\theta \in E'$.

Proof. (b) $\Rightarrow$ (a) if $||\theta|| \leq \varepsilon$, $\mathcal{L}(\mu)(\theta) = \int e^{\langle\theta,x\rangle} d\mu(x) \leq \int e^{\varepsilon\|x\|} d\mu(x) < +\infty$. Then $B(0,\varepsilon) \subset \operatorname{dom}\mathcal{L}(\mu)$.

(a) $\Rightarrow$ (c) If $B(0,\varepsilon) \subset \operatorname{dom}\mathcal{L}(\mu)$, then $\mathcal{L}(\mu\theta^{-1})(\lambda) < +\infty$ for any λ such that $|\lambda| < \frac{\varepsilon}{\|\theta\|}$.

(c) $\Rightarrow$ (b). Let $(\theta_1, \dots, \theta_d)$ be a base of $E' = E$ and $\Theta = \{\theta_1, \dots, \theta_d, -\theta_1, \dots, -\theta_d\}$. Then $|x| = \sup_{\theta\in\Theta}\langle\theta, x\rangle$ defines a norm equivalent to $||.||$. Therefore $||x|| \leq C|x|$, $e^{\varepsilon\|x\|} \leq \sum_{\theta\in\Theta} e^{C\varepsilon\langle\theta,x\rangle}$ and

$$\int e^{\varepsilon\|x\|} d\mu(x) \leq \sum_{\theta\in\Theta} \mathcal{L}(\mu)(C\varepsilon\theta).$$

Since Θ is finite and $0 \in \operatorname{int}\operatorname{dom}\mathcal{L}(\mu\theta^{-1})$ for any $\theta \in \Theta$, there exists $\varepsilon > 0$ such that the right hand side of the previous inequality is finite. □

Lemma 24. *For any $\mu \in \mathcal{CM}(E)$, $\operatorname{dom}\mathcal{L}(\mu)$ is open and then $\mathcal{L}$ and $\mathcal{C}$ are injective in $\mathcal{CM}(E)$.*

Proof. Since $\mu \in \mathcal{CM}(E)$ and $\theta \in \operatorname{dom}\mathcal{L}(\mu)$ implies $e^{\langle\theta,x\rangle}\mu \in \mathcal{CM}(E)$, we only have to prove $0 \in \operatorname{int}\operatorname{dom}\mathcal{L}(\mu)$ for any $\mu \in \mathcal{CM}(E)$. Using lemma 23 and the fact that $\mu\theta^{-1} \in \mathcal{CM}(\mathbb{R})$ for any $\theta \in E'$ we are reduced to the one dimensional case.

Let $\mu \in \mathcal{CM}(\mathbb{R})$. If μ is a Dirac measure then $\operatorname{dom}\mathcal{L}(\mu) = \mathbb{R}$ is open. Otherwise μ has a density e^{-c} with c convex. From $\mu(\mathbb{R}) = \int e^{-c(x)} dx < +\infty$, we see that c strictly increases somewhere. By the convexity of c, this implies that $c(x) - \varepsilon x$ increases for x large enough and $\varepsilon > 0$. Then, $\sup_{x\geq 0} \varepsilon x - c(x) < +\infty$ and by symmetry we obtain, for some $\varepsilon > 0$, $\sup_{x\in\mathbb{R}} \varepsilon|x| - c(x) < +\infty$ and $\mu(e^{-\frac{\varepsilon|x|}{2}}) < +\infty$ which by lemma 23 allows to conclude.

The openness of $\operatorname{dom}\mathcal{L}(\mu) \ni 0$ implies that μ is characterized by its Laplace transform $\mathcal{L}(\mu)$. Since $\log\mathcal{L}(\mu)$ is a l.s.c. convex function, it is also characterized by its Fenchel transform $\mathcal{C}(\mu)$. Then $\mathcal{L}$ and $\mathcal{C}$ are injective on $\mathcal{CM}(E)$. □

Proof of theorem 12. Let us now prove the tightness of $\mathcal{C}(\mu_n)$ when $\mu_n \xrightarrow{\text{w}} \mu$. From the previous results we know that there exists $\varepsilon > 0$ such that $\int e^{\langle\theta,x\rangle} d\mu(x) < +\infty$ for any θ such that $||\theta|| \leq \varepsilon$. Let us choose a finite set $\Theta \subset B(0,\varepsilon)$ such that $||x|| \leq \frac{1}{\varepsilon}\sup_{\theta\in\Theta}\langle\theta, x\rangle$. Since (by the proof of proposition 21) $\limsup_n \mathcal{L}(\mu_n)(\theta) \le \mathcal{L}(\mu)(\theta)$ for any $\theta \in E'$, we have $\mathcal{L}(\mu_n)(\theta) \le C_0$ for $n \ge n_0$ and $\theta \in \Theta$. Then,

$$\begin{aligned}\mathcal{C}(\mu_n)(x) &= \sup_{\theta\in E'} \langle\theta, x\rangle - \log \mathcal{L}(\mu_n)(\theta) \\ &\ge \sup_{\theta\in\Theta}\langle\theta, x\rangle - C_0 \ge \varepsilon\|x\| - C_0.\end{aligned}$$

This shows that the sequence $\mathcal{C}(\mu_n)$ for $n \ge n_0$ is tight and by the previous section and the plan of section 2 we obtain the weak convergence of $\mathcal{C}(\mu_n)$ towards $\mathcal{C}(\mu)$.

By Lemma 24, the Cramer transform is injective and we can consider the continuity of the inverse map $\mathcal{C}^{-1}$ on the image of $\mathcal{CM}(E)$. We only have to prove that for any μ_n and $\mu \in \mathcal{CM}(E)$, $\mathcal{C}(\mu_n) \overset{\mathrm{w}}{\to} \mathcal{C}(\mu)$ implies $\mu_n \overset{\mathrm{w}}{\to} \mu$. But since any sequence of finite measures on $\mathbb{R}^d$ admits a vaguely converging subsequence, we have $\mu_n \overset{\mathrm{v}}{\to} \nu$ for a subsequence of μ_n also denoted μ_n. This implies the logconcavity of ν and then this is equivalent to the weak convergence and implies $\mathcal{C}(\mu_n) \overset{\mathrm{w}}{\to} \mathcal{C}(\nu)$. Then $\mathcal{C}(\nu) = \mathcal{C}(\mu)$ and by the injectivity of $\mathcal{C}$, we get $\nu = \mu$. By the unicity of the limit we obtain $\mu_n \overset{\mathrm{w}}{\to} \nu$. $\square$

5 The infinite dimensional case

From Theorem 12 (proved in sections 3 and 4) and the plan given in section 2, the proof of Theorem 13 is reduced to the proof of the tightness of $\mathcal{C}(\mu_n)$ when $\mu_n \overset{\mathrm{w}}{\to} \mu$ and $\mu_n,\ \mu \in \mathcal{CM}(E)$. Since $\mu_n \overset{\mathrm{w}}{\to} \mu$ and E is a Polish space, the set composed by the measures μ_n and μ is tight. Moreover there exists $0 < \alpha < \beta < +\infty$ such that $\alpha \le \nu(E) \le \beta$ for any measure ν in this set. We say that a set of positive measures is bounded when it satisfies this last condition. The following result concludes the proof of Theorem 13.

Proposition 25. *Let E and X be as in Theorem* 13*. The image by the Cramer transform of a tight and bounded subset of $\mathcal{CM}(E)$ is a tight subset of $\mathcal{C}(E)$.*

Proof. Let Π be a tight and bounded subset of $\mathcal{CM}(E)$. There exists a compact subset C of E such that

$$\mu(C^c) \le \frac{1}{4}\mu(E) \quad \text{for any } \mu \in \Pi. \tag{6}$$

Let us consider the closed convex and symmetric (stable by $x \mapsto -x$) hull of C. Since E is convex and complete, the hull of C is still compact and satisfies (6). We denote it again by C. Let φ be its support function, $\varphi(\theta) = \sup_{x\in C}\langle\theta, x\rangle$. From the symmetry of C, φ is a seminorm on X' and we denote by $B' = \{\theta \in X',\ \varphi(\theta) \le 1\}$ its unit ball. Since C is a closed convex set, $\mathbb{1}_C(x) = \mathcal{F}(\varphi)(x) = \sup_{\theta\in X'}\langle\theta, x\rangle - \varphi(\theta)$. Then $x \in C$ iff $\langle\theta, x\rangle \le 1$ for any $\theta \in B'$ and the gauge function of C is $J_C(x) \overset{\mathrm{def}}{=} \inf_{\lambda,\, x\in\lambda C} \lambda = \sup_{\theta\in B'}\langle\theta, x\rangle$.

Let us now upper bound $\mathcal{L}(\mu)$ for any $\mu \in \Pi$. For any $\theta \in X'$, $\nu = \mu\theta^{-1}$ is logconcave on $\mathbb{R}$ and $\mathcal{L}(\mu)(\lambda\theta) = \mathcal{L}(\nu)(\lambda)$. Let us fix $\theta \in B'$. Then $\theta^{-1}([-1, 1]) \supset C$ and $\nu([-1, 1]^c) \le \frac{1}{4}\nu(\mathbb{R})$.

However, by the logconcavity of ν, we have for $n \geq 1$

$$\nu((1,3]) \geq \nu([-1,1])^{1-\frac{1}{n}} \nu((-1+2n, 1+2n])^{\frac{1}{n}}$$

then

$$\nu((-1+2n, 1+2n]) \leq \frac{\nu(\mathbb{R})}{3^n}.$$

Therefore if $e^{2\lambda} < 3$,

$$\begin{aligned}
\mathcal{L}(\mu)(\lambda\theta) &= \mathcal{L}(\nu)(\lambda) = \int_{\mathbb{R}} e^{\lambda x} \nu(dx) \\
&\leq e^{\lambda} \nu((-\infty, 1]) + \sum_{n=1}^{\infty} e^{\lambda(1+2n)} \nu((-1+2n, 1+2n]) \\
&\leq \frac{e^{\lambda}}{1 - \frac{e^{2\lambda}}{3}} \mu(E).
\end{aligned}$$

Then, for some constant a independent of Π, $\varphi(\theta) \leq a$ implies $\log \mathcal{L}(\mu)(\theta) \leq \log \mu(E) + 1$. If β is an upper bound of $\mu(E)$ for $\mu \in \Pi$, we obtain $\mathcal{C}(\mu)(x) = \sup_{\theta \in X'} \langle \theta, x \rangle - \log \mathcal{L}(\mu)(\theta) \geq a J_C(x) - \log \beta - 1$. Then $\inf_{\mu \in \Pi} \inf_{x \in (kC)^c} \mathcal{C}(\mu)(x) \geq ak - \log \beta - 1$ which by the compacity of sets kC implies the tightness of $\mathcal{C}(\Pi)$. □

A Proof of Proposition 11

As was said in section 2, one implication is a consequence of theorems 1 and 2 of [24]. Indeed, if the support of μ, $\operatorname{Supp} \mu \subset F$ and F is an affine space, the logconcavity of μ is equivalent to that of $\mu|_F$ and if F has zero dimension, μ is obviously logconcave.

Let μ be a logconcave measure on E. If we denote by $B(x, \varepsilon) = x + \varepsilon B(0,1)$ the open ball of center x and radius ε for the euclidian norm, we have $(1-t)B(x,\varepsilon) + tB(y,\varepsilon) = B((1-t)x + ty, \varepsilon)$. Then, the support of μ is convex. Indeed, if x and $y \in \operatorname{Supp} \mu$, for any $\varepsilon > 0$, $\mu(B(x,\varepsilon)) > 0$ and $\mu(B(y,\varepsilon)) > 0$ and by the logconcavity of μ, $\mu(B((1-t)x + ty, \varepsilon)) > 0$ for any $t \in [0,1]$. Let F be the affine hull of $\operatorname{Supp} \mu$, $\mu|_F$ is logconcave and we may then suppose without restriction that $F = E$. Then, by the convexity of $\operatorname{Supp} \mu$, its interior is non empty [26]. If E has zero dimension, μ is a Dirac measure and we have nothing to prove.

Let us first consider the one dimensional case. Since $C = (-\infty, 0]$ is convex, $f(x) = \mu(C+x) = \mu((-\infty, x])$ is a logconcave function on $\mathbb{R}$ and $\operatorname{dom} f \stackrel{\text{def}}{=} \{x \in \mathbb{R}, f(x) \neq 0\}$ is the classical domain of the convex function $-\log f$ and it contains the interior of $\operatorname{Supp} \mu$. Indeed, if $x \in \operatorname{int} \operatorname{Supp} \mu$, $x - \varepsilon \in \operatorname{Supp} \mu$ for ε small enough, then $\mu(C + x) \geq \mu((x - 2\varepsilon, x)) > 0$. Moreover, since f is a nondecreasing function $\operatorname{dom} f \supset \operatorname{int} \operatorname{Supp} \mu + [0, +\infty)$. By the convexity of $-\log f$, f is continuous in the interior of its domain, then in $\operatorname{int} \operatorname{Supp} \mu + [0, +\infty)$. By symmetry, we obtain that f is continuous on $\mathbb{R}$. Indeed, $g(x) = \mu((x, +\infty)) = \mu(\mathbb{R}) - f(x)$ is nonincreasing, logconcave and its domain contains $\operatorname{int}\operatorname{Supp}\mu$ and then $\operatorname{dom} g \supset \operatorname{int}\operatorname{Supp}\mu + (-\infty, 0]$. This implies that $\mu(\{a\}) = 0$ for any $a \in \mathbb{R}$ and then $\mu(\partial \operatorname{Supp}\mu) = 0$ and also $\operatorname{dom} f = \operatorname{int}\operatorname{Supp}\mu + [0, +\infty)$.

Since $-\log f$ is convex, it admits finite right and left derivatives in each point of its open domain $\operatorname{int}\operatorname{Supp}\mu + [0, +\infty)$ and so does f. Since the complementary of this domain is stable by negative translation and $f \equiv 0$ there, the left derivative of f exists and is null. Then, $f'_-(x) = \lim_{\varepsilon \to 0_+} \frac{\mu(x+(-\varepsilon,0])}{\varepsilon}$ exists and is finite for any $x \in \mathbb{R}$. Moreover since $(-\varepsilon, 0]$ is convex, f'_- is again a logconcave function and then it is continuous in the interior of its domain which contains $\operatorname{int}\operatorname{Supp}\mu$. Indeed, if $x \in \operatorname{int}\operatorname{Supp}\mu$, then $f(x) > 0$ and if in addition $f'_-(x) = 0$, we obtain $(\log f)'_-(x) = 0$. Since $\log f$ is a concave and nondecreasing function, this implies $f(y) = f(x)$ for any $y \geq x$ and then $\mu(x + (0, +\infty)) = 0$ which leads to a contradiction with $x \in \operatorname{int}\operatorname{Supp}\mu$.

We have then proved that $g = f'_-$ exists, is finite and is logconcave on $\mathbb{R}$ and that it is strictly positive and continuous and then equal to f' in $\operatorname{int}\operatorname{Supp}\mu$. Moreover g is null on $(\operatorname{Supp}\mu)^c$ and $\mu((\operatorname{int}\operatorname{Supp}\mu)^c) = 0$. This implies that g is a density of μ on $\mathbb{R}$.

We now solve the n dimensional case by induction. Let us first note that the previous proof is still valid if μ is not a measure on $\mathbb{R}$ but only a bounded, positive, logconvave (satisfying (1)), additive and nondecreasing functional on convex sets (that is such that $\mu(C \cup C') = \mu(C) + \mu(C')$ if $C \cap C' = \emptyset$ and $\mu(C) \leq \mu(C')$ if $C \subset C'$) with a support with nonempty interior (the support is defined as for measures, that is $x \in \operatorname{Supp}\mu$ iff $\mu(B(x, \varepsilon)) > 0$ for all $\varepsilon > 0$). Indeed the logconcavity implies the continuity of f without assuming the continuity of μ and the continuity of f implies the continuity of μ on convex sets (intervals). For instance if $C_n = (a_n, b_n] \searrow \emptyset$ and is neither empty, we have $a_n \nearrow a$ and $b_n \searrow a$ and then $\mu(C_n) = f(b_n) - f(a_n) \searrow 0$. The conclusion of the proof is at least that μ has a density g which is logconcave and finite everywhere in the sense that $\mu(I) = \int_I g(x)dx$ for any interval I.

Let us consider a bounded, positive, logconcave, additive and nondecreasing functional μ on the convex sets of $\mathbb{R}^n$, with a support with nonempty interior, and suppose that any functional of this type on $\mathbb{R}^{n-1}$ has a density g which is finite and logconcave everywhere in the sense that $\mu(C) = \int_C g(x)dx$ for any rectangular boxes $C = I_1 \times \ldots I_{n-1}$ (where I_j are intervals of $\mathbb{R}$) and that

$$g(x) = \lim_{\varepsilon_{n-1} \to 0_+} \ldots \lim_{\varepsilon_1 \to 0_+} \frac{\mu(x + (-\varepsilon_1, 0] \times \cdots \times (-\varepsilon_{n-1}, 0])}{\varepsilon_1 \ldots \varepsilon_{n-1}}.$$

For any convex set C of $\mathbb{R}^{n-1}$ the functional $\mu_C : I \mapsto \mu(I \times C)$ is a bounded, positive, logconcave, additive and nondecreasing functional on the convex sets of $\mathbb{R}$. If $\operatorname{Supp}\mu_C$ has a nonempty interior, then μ_C has a density, that is $\mu(I \times C) = \int_I g(x_1, C)dx_1$ for any interval I where

$$g(x_1, C) = \lim_{\varepsilon_1 \to 0_+} \frac{\mu((x_1 + (-\varepsilon_1, 0]) \times C)}{\varepsilon_1}$$

is finite for any $x_1 \in \mathbb{R}$. If $\mu_C \equiv 0$ this is again true. If now $\operatorname{Supp}\mu_C$ is non empty with empty interior, it is then equal to some singleton $\{a\}$. Therefore, $\operatorname{Supp}\mu \subset \{a\} \times C \cup \mathbb{R} \times (\operatorname{int} C)^c$. The convexity of $\operatorname{Supp}\mu$ and the fact that it has non empty interior implies that $\operatorname{Supp}\mu \subset \mathbb{R} \times (\operatorname{int} C)^c$ and then $\mu(\{a\} \times \partial C) = \mu_C(\{a\}) > 0$. If now C' is a convex set such that $\operatorname{int} C' \supset \overline{C}$, either $\mu_{C'}$ has a support with empty interior and then $\mu(\{a\} \times \partial C) \leq \mu(\mathbb{R} \times \operatorname{int} C') = 0$, either $\mu_{C'}$ has a density and then $\mu(\{a\} \times \partial C) \leq \mu_{C'}(\{a\}) = 0$. The two cases contradict the previous conclusion, then all convex sets C are such that $\operatorname{Supp} \mu_C$ is empty or has nonempty interior and then such that $g(x_1, C)$ exists and is a density of μ_C.

For any $x_1 \in \mathbb{R}$, $\mu^{x_1} : C \mapsto g(x_1, C)$ is a bounded (by $g(x_1, \mathbb{R}^{n-1})$ which is finite), positive, logconcave, additive and nondecreasing functional on the convex sets of $\mathbb{R}^{n-1}$. If $\operatorname{Supp} \mu^{x_1}$ has non empty interior then μ^{x_1} has a density, that is $g(x_1, C) = \int_C h(x_1, x_2, \ldots, x_n) dx_2 \ldots dx_n$ for any rectangular boxes C, with

$$h(x) = \lim_{\varepsilon_n \to 0_+} \cdots \lim_{\varepsilon_1 \to 0_+} \frac{\mu(x + (-\varepsilon_1, 0] \times \cdots \times (-\varepsilon_n, 0])}{\varepsilon_1 \ldots \varepsilon_n}.$$

If μ^{x_1} is null, this is also true. If now $\operatorname{Supp} \mu^{x_1}$ is nonempty with empty interior, it is then a subset of a $n-2$ dimensional affine subspace F of $\mathbb{R}^{n-1}$. By rotation and translation, we can suppose that $F = \{0\} \times \mathbb{R}^{n-2}$. Then $g(x_1, \{0\} \times \mathbb{R}^{n-2}) > 0$ and then $\mu((x_1 + (-\varepsilon, 0]) \times \{0\} \times \mathbb{R}^{n-2}) > 0$ for $\varepsilon > 0$ small enough. Exchanging x_1 and x_2 coordinates in the previous reasoning, we see that $C \mapsto \mu((x_1 + (-\varepsilon, 0]) \times C \times \mathbb{R}^{n-2})$ has necessarily a density and then cannot be strictly positive on a singleton which contradicts the assumption on μ^{x_1}. Therefore, μ^{x_1} has always $h(x_1, .)$ as density and μ has h as density (on rectangular boxes). The induction is then proved and if μ is a measure, h is also a density of μ on Borel sets. Moreover, since h is continuous in the interior and exterior of $\operatorname{Supp} \mu$ and $\mu(\partial \operatorname{Supp} \mu) = 0$, we can repace h by its u.s.c. envelope. $\square$

References

[1] M. Akian, *Densities of idempotent measures and large deviations*, Rapport de Recherche 2534, INRIA, 1995.

[2] ______, *Theory of cost measures : convergence of decision variables*, Rapport de Recherche 2611, INRIA, 1995.

[3] ______, *From probabilities to cost measures via large deviations*, Rapport de recherche, INRIA, 1996, to appear.

[4] M. Akian, J. P. Quadrat, and M. Viot, *Bellman processes*, 11th International Conference on Analysis and Optimization of Systems : Discrete Event Systems, Lecture notes in Control and Information Sciences, vol. 199, Springer Verlag, 1994.

[5] ______, *Duality between probability and optimization*, Idempotency (J. Gunawardena, ed.), Cambridge University Press, 1996, to appear.

[6] H. Attouch, *Variational convergence for functions and operators*, Pitman, 1984.

[7] R. Azencott, Y. Guivarc'h, and R. F. Gundy, *Ecole d'été de probabilités de Saint Flour VIII-1978*, Lecture Notes in Mathematics, vol. 774, Springer Verlag, Berlin, 1980.

[8] R. R. Bahadur and S. L. Zabell, *Large deviations of the sample mean in general vector spaces*, Annals of Probab. **7** (1979), 587–621.

[9] F. Bellalouna, *Un point de vue linéaire sur la programmation dynamique. détection de ruptures dans le cadre des problèmes de fiabilité*, Thèse, Université Paris-IX Dauphine, Paris, 1992.

[10] H. J. Brascamp and E. H. Lieb, *On extensions of the Brunn-Minkowski and Prépoka-Leinder theorems, including inequalities for log concave functions, and with an application to the diffusion equation*, Journal of functional analysis **22** (1976), 366–389.

[11] H. Cramer, *Sur un nouveau théorème limite de la théorie des probabilités*, Actualités Sci. Indust. **736** (1938), 5–23.

[12] P. Del Moral, *Résolution particulaire des problèmes d'estimation et d'optimisation non-linéaires*, Thèse, Université Paul Sabatier, Toulouse, 1994.

[13] J. D. Deuschel and D. W. Stroock, *Large deviations*, Pure and applied Mathematics, vol. 137, Academic Press, London, 1989.

[14] R. S. Ellis, *Entropy, large deviations, and statistical mechanics*, Springer Verlag, New York, 1985.

[15] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, Springer Verlag, Berlin, 1984.

[16] J. L. Joly, *Une famille de topologies sur l'ensemble des fonctions convexes pour lesquelles la polarité est bicontinue*, J. Math. pures et appl. **52** (1973), 421–441.

[17] V. N. Kolokoltsov and V. P. Maslov, *The general form of the endomorphisms in the space of continuous functions with values in a numerical commutative semiring (with the operation $\oplus = \max$)*, Soviet Math. Dokl. **36** (1988), no. 1, 55–59.

[18] V. Maslov, *Méthodes opératorielles*, Éditions MIR, Moscou, 1987.

[19] R. Michel, *Logconcavité, équation de la chaleur et viscosité*, International journal of Mathematics **3** (1992), no. 4, 499–524.

[20] R. Michel and M. Volle, *Logconcavité et inégalités intégrales*, Revue Roumaine de Mathématiques Pures et Appliquées **38** (1993), no. 6, 513–530.

[21] G. L. O'Brien, *Sequences of capacities, with connections to large-deviation theory*, J. theoretical probab. (1996).

[22] G. L. O'Brien and W. Vervaat, *Capacities, large deviations and loglog laws*, Stable processes and related topics (S. Cambanis, G. Samorodnitsky, and M.S. Taqqu, eds.), Progress in probability, vol. 25, Birkhaüser, 1991, pp. 43–83.

[23] A. Prékopa, *Logarithmic concave measures with application to stochastic programming*, Acta Sci. Math. **32** (1971), 301–315.

[24] ______, *On logarithmic concave measures and functions*, Acta Sci. Math. **34** (1972), 335–343.

[25] J. P. Quadrat, *Théorèmes asymptotiques en programmation dynamique*, Note CRAS **311** (1990), 745–748.

[26] R. T. Rockafellar, *Convex analysis*, Princeton University Press Princeton, N.J., 1970.

[27] S. R. S. Varadhan, *Large deviations and applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 46, SIAM, Philadelphia, Penn., 1984.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399